

On uniqueness theorems for Tsallis entropy and Tsallis relative entropy

Shigeru Furuichi, *Member, IEEE*,

Abstract—The uniqueness theorem for Tsallis entropy was presented in H. Suyari, *IEEE Trans. Inform. Theory*, Vol.50, pp.1783-1787 (2004) by introducing the generalized Shannon-Khinchin's axiom. In the present paper, this result is generalized and simplified as follows: *Generalization*: The uniqueness theorem for Tsallis relative entropy is shown by means of the generalized Hobson's axiom. *Simplification*: The uniqueness theorem for Tsallis entropy is shown by means of the generalized Faddeev's axiom.

Index Terms—generalized Faddeev's axiom, generalized Hobson's axiom, generalized Shannon-Khinchin's axiom, Tsallis entropy, Tsallis relative entropy, uniqueness theorem.

I. INTRODUCTION

Shannon entropy is uniquely determined by the Shannon-Khinchin's axiom [1], which is referred to as the uniqueness theorem for Shannon entropy. The Shannon-Khinchin's axiom was improved by A. D. Faddeev [2] in the sense that the conditions of his axiom are simpler than those of the Shannon-Khinchin's axiom. (For details, see [3], [4], [5], [6].) As a generalization of the axiomatic characterization of relative entropy, the uniqueness theorem for relative entropy was proven by A. Hobson [7]. Moreover, recently, the nonextensive entropies, including Tsallis entropy, were characterized by H. Suyari in terms of the generalized Shannon-Khinchin's axiom [8]. The uniqueness theorem obtained by a generalization of the Shannon-Khinchin's axiom for the structural α -entropy, which is one of the nonextensive entropies, was the first appearance of such generalized results [9].

The present paper proves the uniqueness theorem for Tsallis relative entropy by combining the axioms of Hobson and Suyari. This uniqueness theorem is a simultaneous generalization of their results with respect to the following two points. The Hobson's theorem is generalized as a parametric extension, and the Suyari's theorem is generalized in the sense that the relative entropy function is an extension of the entropy function. In addition, the uniqueness theorem for Tsallis entropy is proven by the simplification of the generalized Shannon-Khinchin's axiom in the sense that the Faddeev's axiom is simpler than Shannon-Khinchin's axiom.

This paper is organized as follows. In Section II, a brief review of Tsallis entropy and Tsallis relative entropy is pre-

sented. In Section III, the generalized Shannon-Khinchin's axiom introduced by H. Suyari and the uniqueness theorem for Tsallis entropy are reviewed. In Section IV, the uniqueness theorem for Tsallis relative entropy by means of the generalized Hobson's axiom is shown, and the function $\phi(q)$, which will be described in Theorem IV.3, is characterized. In Section V, the uniqueness theorem for Tsallis entropy is shown by means of the generalized Faddeev's axiom. Finally, a theorem on the relation among the generalized Shannon-Khinchin's axiom, the generalized Faddeev's axiom, and Tsallis entropy is presented.

II. TSALLIS ENTROPY AND TSALLIS RELATIVE ENTROPY

A. Tsallis entropy

Several extensions of entropy have been formulated and investigated [10], [11], [12], [13], [14], [15]. Prevalent among these extensions is the Rényi entropy [10], [11], [13]:

$$R_q(X) = \frac{1}{1-q} \log \sum_{i=1}^n p_i^q, \quad (q \neq 1), \quad (1)$$

which has the additivity property

$$R_q(X \times Y) = R_q(X) + R_q(Y), \quad (2)$$

for two independent random variables X and Y . Tsallis entropy introduced in [14], the definition of which is presented herein, the structural α -entropy [9], and the entropy of type β introduced in [12] are referred to as nonadditive (nonextensive) entropies because these entropies do not have an additivity for two independent random variables, whereas the Rényi entropy introduced in [10] and Shannon entropy are referred to as additive (extensive) entropies, due to Eq.(2). The present study examines the nonextensive entropies, including Tsallis entropy as a typical example. Tsallis entropy was defined by C. Tsallis in the field of the statistical physics for the purpose of analyzing multifractal systems [14].

Definition II.1 (Tsallis [14]) For any nonnegative real number q and the probability distribution $p_i \equiv p(X = i)$, ($i = 1, \dots, n$) of a given random variable X , Tsallis entropy is defined as follows:

$$S_q(X) = - \sum_{i=1}^n p_i^q \ln_q p_i, \quad (3)$$

with parameter q as an extension of Shannon entropy, where the q -logarithm is defined as $\ln_q(x) \equiv \frac{x^{1-q}-1}{1-q}$ for any non-negative real numbers q and x , and the convention $0 \ln_q(\cdot) \equiv 0$ is set.

The Tsallis entropy $S_q(X)$ converges to Shannon entropy $-\sum_{i=1}^n p_i \log p_i$ as $q \rightarrow 1$, because the q -logarithm uniformly converges to a natural logarithm as $q \rightarrow 1$. Tsallis entropy plays an essential role in nonextensive statistics, which is often called Tsallis statistics, so that several important findings have been published [15]. In addition, since the q -logarithm function $\ln_q(x)$ has the pseudoadditivity property for $q \neq 1$:

$$\ln_q(xy) = \ln_q(x) + \ln_q(y) + (1-q)\ln_q(x)\ln_q(y), \quad (4)$$

Tsallis entropy has the pseudoadditivity property for $q \neq 1$:

$$S_q(X \times Y) = S_q(X) + S_q(Y) + (1-q)S_q(X)S_q(Y). \quad (5)$$

Remark II.2 Although a simple transformation exists between the Tsallis entropy and the Rényi entropy, such that

$$R_q(X) = \frac{1}{1-q} \log \{1 + (1-q)S_q(X)\}, \quad (6)$$

their structures differ with respect to their additivities (cf. Eq.(2) and Eq.(5)).

B. Tsallis relative entropy

A relative entropy based on Tsallis entropy which is a typical nonextensive entropy, was formulated and discussed in [16], [17], [18] from a physical point of view. The fundamental properties of Tsallis relative entropy in both classical and quantum systems were investigated in [19] from a mathematical point of view.

Definition II.3 Here, probability distributions a_j and b_j are assumed to satisfy $a_j \geq 0$, $b_j \geq 0$ and $\sum_{j=1}^n a_j = \sum_{j=1}^n b_j = 1$. Tsallis relative entropy between $A = \{a_j\}$ and $B = \{b_j\}$, for any $q \geq 0$, is then defined as

$$D_q(A|B) \equiv -\sum_{j=1}^n a_j \ln_q \frac{b_j}{a_j}. \quad (7)$$

Note that $\lim_{q \rightarrow 1} D_q(A|B) = D_1(A|B) \equiv \sum_{j=1}^n a_j \log \frac{a_j}{b_j}$, which is known as relative entropy (often referred to as Kullback-Leibler information, divergence or cross entropy). For Tsallis relative entropy, several fundamental properties, which are listed below, hold as parametric extensions of relative entropy (for example, [19]).

Proposition II.4 (1) (Continuity) $D_q(A|B)$ is a continuous function for a_j and b_j .
 (2) (Nonnegativity) $D_q(A|B) \geq 0$.
 (3) (Symmetry)

$$D_q(a_{\pi(1)}, \dots, a_{\pi(n)} | b_{\pi(1)}, \dots, b_{\pi(n)}) = D_q(a_1, \dots, a_n | b_1, \dots, b_n).$$

(4) (Possibility of extension)

$$D_q(a_1, \dots, a_n, 0 | b_1, \dots, b_n, 0) = D_q(a_1, \dots, a_n | b_1, \dots, b_n).$$

(5) (Pseudoadditivity)

$$\begin{aligned} D_q(A^{(1)} \times A^{(2)} | B^{(1)} \times B^{(2)}) \\ = D_q(A^{(1)} | B^{(1)}) + D_q(A^{(2)} | B^{(2)}) \\ + (q-1) D_q(A^{(1)} | B^{(1)}) D_q(A^{(2)} | B^{(2)}), \end{aligned}$$

where

$$A^{(1)} \times A^{(2)} = \left\{ a_j^{(1)} a_j^{(2)} \mid a_j^{(i)} \in A^{(i)}, i = 1, 2 \right\},$$

and

$$B^{(1)} \times B^{(2)} = \left\{ b_j^{(1)} b_j^{(2)} \mid b_j^{(i)} \in B^{(i)}, i = 1, 2 \right\}.$$

(6) (Joint convexity) For $0 \leq \lambda \leq 1$, any $q \geq 0$ and the probability distributions $A^{(i)} = \{a_j^{(i)}\}$, $B^{(i)} = \{b_j^{(i)}\}$, ($i = 1, 2$), we have

$$\begin{aligned} D_q(\lambda A^{(1)} + (1-\lambda)A^{(2)} | \lambda B^{(1)} + (1-\lambda)B^{(2)}) \\ \leq \lambda D_q(A^{(1)} | B^{(1)}) + (1-\lambda) D_q(A^{(2)} | B^{(2)}). \end{aligned}$$

(7) (Additivity)

$$\begin{aligned} D_q(a_1, \dots, a_{i-1}, a_{i1}, a_{i2}, a_{i+1}, \dots, a_n | \\ b_1, \dots, b_{i-1}, b_{i1}, b_{i2}, b_{i+1}, \dots, b_n) \\ = D_q(a_1, \dots, a_n | b_1, \dots, b_n) \\ + a_i^q b_i^{1-q} D_q\left(\frac{a_{i1}}{a_i}, \frac{a_{i2}}{a_i} \mid \frac{b_{i1}}{b_i}, \frac{b_{i2}}{b_i}\right), \end{aligned}$$

where $a_i = a_{i1} + a_{i2}$, $b_i = b_{i1} + b_{i2}$.

(8) (Monotonicity) For a transition probability matrix W , we have

$$D_q(WA | WB) \leq D_q(A | B).$$

Conversely, in the present paper, Tsallis relative entropy is axiomatically characterized by the generalized Hobson's axiom.

III. REVIEW OF THE GENERALIZED SHANNON-KHINCHIN'S AXIOMS AND THE UNIQUENESS THEOREM FOR TSALLIS ENTROPY

The generalized Shannon-Khinchin's axiom introduced by H. Suyari is reviewed in the following. The function $S_q(x_1, \dots, x_n)$ is assumed to be defined for the n -tuple (x_1, \dots, x_n) belonging to

$$\Delta_n \equiv \left\{ (p_1, \dots, p_n) \mid \sum_{i=1}^n p_i = 1, p_i \geq 0 \ (i = 1, \dots, n) \right\}$$

and to take a value in $\mathbb{R}^+ \equiv [0, \infty)$.

Axiom III.1 (Generalized Shannon-Khinchin's axiom)

(G1) *Continuity*: The function

$$S_q : \Delta_n \rightarrow \mathbb{R}^+$$

is continuous.

(G2) *Maximality*:

$$S_q\left(\frac{1}{n}, \dots, \frac{1}{n}\right) = \max \{S_q(x_1, \dots, x_n) : x_i \in \Delta_n\} > 0.$$

(G3) *Generalized Shannon additivity*: For $x_{ij} \geq 0$, $x_i = \sum_{j=1}^{m_i} x_{ij}$, ($i = 1, \dots, n; j = 1, \dots, m_i$),

$$S_q(x_{11}, \dots, x_{nm_n}) = S_q(x_1, \dots, x_n) + \sum_{i=1}^n x_i^q S_q\left(\frac{x_{i1}}{x_i}, \dots, \frac{x_{im_i}}{x_i}\right).$$

(G4) *Expandability*:

$$S_q(x_1, \dots, x_n, 0) = S_q(x_1, \dots, x_n).$$

Note that condition (G4) is altered slightly from [GSK4] of the original axiom [8]. In condition (G2), the strict positivity of the maximum value of Tsallis entropy $S_q(x_1, \dots, x_n)$ is also imposed, in addition to [GSK2] of the original axiom [8]. This adoption excludes a trivial situation that Tsallis entropy is constant zero. Then, the following theorem was shown by H. Suyari [8].

Theorem III.2 (Suyari [8]) Conditions (G1) to (G4) determine the function S_q such that

$$S_q(x_1, \dots, x_n) = \frac{1 - \sum_{i=1}^n x_i^q}{\phi(q)}, \quad (8)$$

where $\phi(q)$ is characterized by the following conditions:

- (i) $\phi(q)$ is continuous and $\phi(q)(q-1) > 0$ for $q \neq 1$.
- (ii) $\lim_{q \rightarrow 1} \phi(q) = 0$ and $\phi(q) \neq 0$ for $q \neq 1$.
- (iii) There exists $(a, b) \subset \mathbb{R}^+$ such that $a < 1 < b$ and $\phi(q)$ is differentiable on $(a, 1)$ and $(1, b)$.
- (iv) There exists a positive constant number k such that $\lim_{q \rightarrow 1} \frac{d\phi(q)}{dq} = \frac{1}{k}$.

IV. AXIOMATIC CHARACTERIZATION OF TSALLIS RELATIVE ENTROPY BY THE GENERALIZED HOBSON'S AXIOM

A. Uniqueness theorem for Tsallis relative entropy

The uniqueness theorem for relative entropy was shown by A. Hobson as follows [7]:

Theorem IV.1 (Hobson [7]) The function $D_1(A|B)$ is assumed to be defined for any two probability distributions $A = \{a_j\}$ and $B = \{b_j\}$ for $j = 1, \dots, n$. If $D_1(A|B)$ satisfies the following conditions, then it is given by the form $k \sum_{j=1}^n a_j \log \frac{a_j}{b_j}$ with a positive constant k .

(H1) *Continuity*: $D_1(A|B)$ is a continuous function of $2n$ variables.

(H2) *Symmetry*:

$$D_1(a_1, \dots, a_j, \dots, a_k, \dots, a_n | b_1, \dots, b_j, \dots, b_k, \dots, b_n) = D_1(a_1, \dots, a_k, \dots, a_j, \dots, a_n | b_1, \dots, b_k, \dots, b_j, \dots, b_n).$$

(H3) *Grouping axiom*:

$$\begin{aligned} & D_1(a_{1,1}, \dots, a_{1,m}, a_{2,1}, \dots, a_{2,m} | b_{1,1}, \dots, b_{1,m}, b_{2,1}, \dots, b_{2,m}) \\ &= D_1(c_1, c_2 | d_1, d_2) \\ &+ c_1 D_1\left(\frac{a_{1,1}}{c_1}, \dots, \frac{a_{1,m}}{c_1} \middle| \frac{b_{1,1}}{d_1}, \dots, \frac{b_{1,m}}{d_1}\right) \\ &+ c_2 D_1\left(\frac{a_{2,1}}{c_2}, \dots, \frac{a_{2,m}}{c_2} \middle| \frac{b_{2,1}}{d_2}, \dots, \frac{b_{2,m}}{d_2}\right) \end{aligned}$$

where $c_i = \sum_{j=1}^m a_{i,j}$ and $d_i = \sum_{j=1}^m b_{i,j}$.

(H4) $D_1(A|B) = 0$ if $a_j = b_j$ for all j .

(H5) $D_1(\frac{1}{n}, \dots, \frac{1}{n}, 0, \dots, 0 | \frac{1}{n_0}, \dots, \frac{1}{n_0})$ is an increasing function of n_0 and a decreasing function of n , for any integers n and n_0 such that $n_0 \geq n$.

The function D_q is defined for the probability distributions $A = \{a_j\}$ and $B = \{b_j\}$ on a finite probability space with one parameter $q \geq 0$. Tsallis relative entropy is characterized by means of the following triplet of the generalized conditions (R1), (R2) and (R3).

Axiom IV.2 (Generalized Hobson's axiom)

(R1) *Continuity*: $D_q(a_1, \dots, a_n | b_1, \dots, b_n)$ is a continuous function of $2n$ variables.

(R2) *Symmetry*:

$$D_q(a_1, \dots, a_j, \dots, a_k, \dots, a_n | b_1, \dots, b_j, \dots, b_k, \dots, b_n) = D_q(a_1, \dots, a_k, \dots, a_j, \dots, a_n | b_1, \dots, b_k, \dots, b_j, \dots, b_n).$$

(R3) *Generalized additivity*:

$$\begin{aligned} & D_q(a_{1,1}, \dots, a_{1,m}, \dots, a_{n,1}, \dots, a_{n,m} | b_{1,1}, \dots, b_{1,m}, \dots, b_{n,1}, \dots, b_{n,m}) \\ &= D_q(c_1, \dots, c_n | d_1, \dots, d_n) \\ &+ \sum_{i=1}^n c_i^q d_i^{1-q} D_q\left(\frac{a_{i,1}}{c_i}, \dots, \frac{a_{i,m}}{c_i} \middle| \frac{b_{i,1}}{d_i}, \dots, \frac{b_{i,m}}{d_i}\right), \quad (9) \end{aligned}$$

where $c_i = \sum_{j=1}^m a_{i,j}$ and $d_i = \sum_{j=1}^m b_{i,j}$.

Then, we have the following theorem:

Theorem IV.3 If conditions (R1), (R2) and (R3) hold, then $D_q(A|B)$ is given in the following form:

$$D_q(A|B) = \frac{1 - \sum_{i=1}^n a_i^q b_i^{1-q}}{\phi(q)} \quad (10)$$

with a certain function $\phi(q)$.

Proof: We prove the theorem by using conditions (R1), (R2) and (R3). First, we define

$$f_q(s, t) \equiv D_q\left(\frac{1}{s}, \dots, \frac{1}{s}, 0, \dots, 0 \middle| \frac{1}{t}, \dots, \frac{1}{t}\right) \quad (11)$$

for any natural numbers s and t such that $t \geq s$. From condition (R2), we have

$$\begin{aligned} & D_q \left(\frac{1}{su}, \dots, \frac{1}{su}, 0, \dots, 0, \dots, \frac{1}{su}, \dots, \frac{1}{su}, \right. \\ & \quad \left. 0, \dots, 0 \middle| \frac{1}{tv}, \dots, \frac{1}{tv} \right) \\ &= D_q \left(\frac{1}{su}, \dots, \frac{1}{su}, \dots, \frac{1}{su}, \dots, \frac{1}{su}, \right. \\ & \quad \left. 0, \dots, 0, \dots, 0, \dots, 0 \middle| \frac{1}{tv}, \dots, \frac{1}{tv} \right). \end{aligned}$$

Also, from condition (R3), we have

$$\begin{aligned} & D_q \left(\frac{1}{su}, \dots, \frac{1}{su}, 0, \dots, 0, \dots, \frac{1}{su}, \dots, \frac{1}{su}, \right. \\ & \quad \left. 0, \dots, 0 \middle| \frac{1}{tv}, \dots, \frac{1}{tv} \right) \\ &= D_q \left(\frac{1}{u}, \dots, \frac{1}{u}, 0, \dots, 0 \middle| \frac{1}{v}, \dots, \frac{1}{v} \right) \\ &+ u \left(\frac{1}{u} \right)^q \left(\frac{1}{v} \right)^{1-q} D_q \left(\frac{1}{s}, \dots, \frac{1}{s}, 0, \dots, 0 \middle| \frac{1}{t}, \dots, \frac{1}{t} \right). \end{aligned}$$

From the above two equations, we have

$$\begin{aligned} & D_q \left(\frac{1}{su}, \dots, \frac{1}{su}, \dots, \frac{1}{su}, \dots, \frac{1}{su}, 0, \dots, \right. \\ & \quad \left. 0, \dots, 0, \dots, 0 \middle| \frac{1}{tv}, \dots, \frac{1}{tv} \right) \\ &= D_q \left(\frac{1}{u}, \dots, \frac{1}{u}, 0, \dots, 0 \middle| \frac{1}{v}, \dots, \frac{1}{v} \right) \\ &+ u \left(\frac{1}{u} \right)^q \left(\frac{1}{v} \right)^{1-q} D_q \left(\frac{1}{s}, \dots, \frac{1}{s}, 0, \dots, 0 \middle| \frac{1}{t}, \dots, \frac{1}{t} \right). \end{aligned}$$

Thus, we have

$$f_q(su, tv) = f_q(u, v) + \left(\frac{u}{v} \right)^{1-q} f_q(s, t).$$

Using a method similar to that used for Lemma 5 in [9], we find that there exists a certain function $\phi(q)$, which depends on the parameter q only, such that

$$f_q(s, t) = \frac{1 - \left(\frac{s}{t} \right)^{1-q}}{\phi(q)}.$$

Here, all a_i and b_i , ($i = 1, \dots, n$) are assumed to be rational numbers. Then, setting $c_i = \frac{l_i}{\sum_{k=1}^n l_k}$, $d_i =$

$\frac{m_i}{\sum_{k=1}^n m_k}$, ($i = 1, \dots, n$) for some natural numbers l_i and m_i such that $l_i \leq m_i$, and we set $a_{i,j} = \frac{1}{\sum_{k=1}^n l_k}$ ($i = 1, \dots, n; j = 1, \dots, l_i$), $b_{i,j} = \frac{1}{\sum_{k=1}^n m_k}$ ($i = 1, \dots, n; j = 1, \dots, m_i$). Substituting into

condition (R3), we have

$$\begin{aligned} & D_q \left(\frac{1}{\sum_{k=1}^n l_k}, \dots, \frac{1}{\sum_{k=1}^n l_k}, 0, \dots, 0, \dots, \frac{1}{\sum_{k=1}^n l_k}, \right. \\ & \quad \left. \dots, \frac{1}{\sum_{k=1}^n l_k}, 0, \dots, 0 \middle| \frac{1}{\sum_{k=1}^n m_k}, \dots, \frac{1}{\sum_{k=1}^n m_k} \right) \\ &= D_q(c_i, \dots, c_n | d_1, \dots, d_n) \\ &+ \sum_{i=1}^n c_i^q d_i^{1-q} D_q \left(\frac{1}{l_i}, \dots, \frac{1}{l_i}, 0, \dots, 0 \middle| \frac{1}{m_i}, \dots, \frac{1}{m_i} \right). \end{aligned}$$

Thus, we have

$$\begin{aligned} & D_q(c_i, \dots, c_n | d_1, \dots, d_n) \\ &= f_q \left(\sum_{k=1}^n l_k, \sum_{k=1}^n m_k \right) - \sum_{i=1}^n c_i^q d_i^{1-q} f_q(l_i, m_i) \\ &= \frac{1 - \left(\sum_{k=1}^n l_k \right)^{1-q} \left(\sum_{k=1}^n m_k \right)^{q-1}}{\phi(q)} \\ &\quad - \frac{\sum_{i=1}^n c_i^q d_i^{1-q} \left(1 - l_i^{1-q} m_i^{q-1} \right)}{\phi(q)} \\ &= \frac{1 - \sum_{i=1}^n c_i^q d_i^{1-q}}{\phi(q)}. \end{aligned}$$

From condition (R1) and the fact that any real number can be approximated by a rational number, the above result holds for any positive real numbers a_i and b_i satisfying $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i = 1$. ■

We give a remark on the conditions of Axiom IV.2 in the following proposition.

Proposition IV.4 The following conditions (R3') and (R4), and the symmetry (R2) imply the condition (R3).

(R3') *Generalized grouping axiom:*

$$\begin{aligned} & D_q(a_{1,1}, \dots, a_{1,m}, a_{2,1}, \dots, a_{2,m} \\ & \quad | b_{1,1}, \dots, b_{1,m}, b_{2,1}, \dots, b_{2,m}) \\ &= D_q(c_1, c_2 | d_1, d_2) \\ &+ c_1^q d_1^{1-q} D_q \left(\frac{a_{1,1}}{c_1}, \dots, \frac{a_{1,m}}{c_1} \middle| \frac{b_{1,1}}{d_1}, \dots, \frac{b_{1,m}}{d_1} \right) \\ &+ c_2^q d_2^{1-q} D_q \left(\frac{a_{2,1}}{c_2}, \dots, \frac{a_{2,m}}{c_2} \middle| \frac{b_{2,1}}{d_2}, \dots, \frac{b_{2,m}}{d_2} \right) \end{aligned}$$

where $c_i = \sum_{j=1}^m a_{i,j}$ and $d_i = \sum_{j=1}^m b_{i,j}$.

(R4) $D_q(A|B) = 0$ if $a_j = b_j$ for all j .

Proof: Here, the notation $\widetilde{D}_q(a_1, \dots, a_n | *) \equiv D_q(a_1, \dots, a_n | b_1, \dots, b_n)$ is introduced in order to simplify

the proof. From condition (R3'), we have

$$\begin{aligned} \widetilde{D}_q(a_1, \dots, a_n | *) &= \widetilde{D}_q(A, a_{n-1} + a_n | *) \\ &+ A^q B^{1-q} \widetilde{D}_q\left(\frac{a_1}{A}, \dots, \frac{a_{n-2}}{A} | *\right) \\ &+ (a_{n-1} + a_n)^q (b_{n-1} + b_n)^{1-q} \\ &\times \widetilde{D}_q\left(\frac{a_{n-1}}{a_{n-1} + a_n}, \frac{a_n}{a_{n-1} + a_n} | *\right) \end{aligned} \quad (12)$$

and

$$\begin{aligned} \widetilde{D}_q(a_1, \dots, a_{n-2}, a_{n-1} + a_n | *) &= \widetilde{D}_q(A, a_{n-1} + a_n | *) \\ &+ A^q B^{1-q} \widetilde{D}_q\left(\frac{a_1}{A}, \dots, \frac{a_{n-2}}{A} | *\right) \\ &+ (a_{n-1} + a_n)^q (b_{n-1} + b_n)^{1-q} D_q(1 | 1) \end{aligned} \quad (13)$$

where $A = a_1 + \dots + a_{n-2}$ and $B = b_1 + \dots + b_{n-2}$. Thus the condition (R4) and the above equations Eq.(12) and Eq.(13) imply

$$\begin{aligned} \widetilde{D}_q(a_1, \dots, a_n | *) &= \widetilde{D}_q(a_1, \dots, a_{n-2}, a_{n-1} + a_n | *) \\ &+ (a_{n-1} + a_n)^q (b_{n-1} + b_n)^{1-q} \\ &\times \widetilde{D}_q\left(\frac{a_{n-1}}{a_{n-1} + a_n}, \frac{a_n}{a_{n-1} + a_n} | *\right) \end{aligned} \quad (14)$$

Next, we derive condition (R3) by induction on n . Assuming that condition (R3) holds for some natural number n , we have

$$\begin{aligned} \widetilde{D}_q(a_{1,1}, \dots, a_{1,m}, \dots, a_{n+1,1}, \dots, a_{n+1,m} | *) &= \widetilde{D}_q(c_1, \dots, c_{n-1}, c_n + c_{n+1} | *) \\ &+ \sum_{i=1}^{n-1} c_i^q d_i^{1-q} \widetilde{D}_q\left(\frac{a_{i,1}}{c_i}, \dots, \frac{a_{i,m}}{c_i} | *\right) \\ &+ (c_n + c_{n+1})^q (d_n + d_{n+1})^{1-q} \\ &\times \widetilde{D}_q\left(\frac{a_{n,1}}{c_n + c_{n+1}}, \dots, \frac{a_{n,m}}{c_n + c_{n+1}}, \frac{a_{n+1,1}}{c_n + c_{n+1}}, \dots, \frac{a_{n+1,m}}{c_n + c_{n+1}} | *\right), \end{aligned} \quad (15)$$

by the use of (R2), (R4) and Eq.(14). By Eq.(14), the first term in the right-hand side of Eq.(15) can be written as

$$\begin{aligned} \widetilde{D}_q(c_1, \dots, c_{n-1}, c_n + c_{n+1} | *) &= \widetilde{D}_q(c_1, \dots, c_n, c_{n+1} | *) \\ &- (c_n + c_{n+1})^q (d_n + d_{n+1})^{1-q} \\ &\times \widetilde{D}_q\left(\frac{c_n}{c_n + c_{n+1}}, \frac{c_{n+1}}{c_n + c_{n+1}} | *\right). \end{aligned} \quad (16)$$

By condition (R3'), the last term in the right-hand side of Eq.(15) can be written as

$$\begin{aligned} (c_n + c_{n+1})^q (d_n + d_{n+1})^{1-q} &\times \widetilde{D}_q\left(\frac{a_{n,1}}{c_n + c_{n+1}}, \dots, \frac{a_{n,m}}{c_n + c_{n+1}}, \frac{a_{n+1,1}}{c_n + c_{n+1}}, \dots, \frac{a_{n+1,m}}{c_n + c_{n+1}} | *\right) \\ &= (c_n + c_{n+1})^q (d_n + d_{n+1})^{1-q} \end{aligned}$$

$$\begin{aligned} &\times \widetilde{D}_q\left(\frac{c_n}{c_n + c_{n+1}}, \frac{c_{n+1}}{c_n + c_{n+1}} | *\right) \\ &+ c_n^q d_n^{1-q} \widetilde{D}_q\left(\frac{a_{n,1}}{c_n}, \dots, \frac{a_{n,m}}{c_n} | *\right) \\ &+ c_{n+1}^q d_{n+1}^{1-q} \widetilde{D}_q\left(\frac{a_{n+1,1}}{c_{n+1}}, \dots, \frac{a_{n+1,m}}{c_{n+1}} | *\right). \end{aligned} \quad (17)$$

Substituting Eq.(16) and Eq.(17) into Eq.(15), we have

$$\begin{aligned} \widetilde{D}_q(a_{1,1}, \dots, a_{1,m}, \dots, a_{n+1,1}, \dots, a_{n+1,m} | *) &= \widetilde{D}_q(c_1, \dots, c_n, c_{n+1} | *) \\ &+ \sum_{i=1}^{n-1} c_i^q d_i^{1-q} \widetilde{D}_q\left(\frac{a_{i,1}}{c_i}, \dots, \frac{a_{i,m}}{c_i} | *\right) \\ &+ c_n^q d_n^{1-q} \widetilde{D}_q\left(\frac{a_{n,1}}{c_n}, \dots, \frac{a_{n,m}}{c_n} | *\right) \\ &+ c_{n+1}^q d_{n+1}^{1-q} \widetilde{D}_q\left(\frac{a_{n+1,1}}{c_{n+1}}, \dots, \frac{a_{n+1,m}}{c_{n+1}} | *\right) \end{aligned}$$

which means that condition (R3) holds for $n + 1$. Thus, the proposition is proven. ■

From Proposition IV.4, we find that we may adopt the axiom composed of the set of $\{(R1), (R2), (R3') \text{ and } (R4)\}$ instead of the set of $\{(R1), (R2) \text{ and } (R3)\}$ in Theorem IV.3.

Condition (R3') of the present paper is a generalization of (H3) in the sense that the factor $c_i^q d_i^{q'}$ of the second term in the right-hand side of Eq.(9) is a parametric extension of the factor $c_i^1 d_i^0$ of that of Eq.(9), satisfying $q + q' = 1$. Condition (R3') of the present study contains the information on d_i as a factor of each function D_q in the second term in the right-hand side of Eq.(9), whereas the original condition (H3) does not, because it is eliminated by $d_i^0 = 1$ in such a special case.

B. Characterization of $\phi(q)$

In this subsection, the function $\phi(q)$ appeared in Theorem IV.3 is characterized.

Proposition IV.5 The property whereby Tsallis relative entropy is one parametric extension of relative entropy:

$$\lim_{q \rightarrow 1} D_q(A|B) = k \sum_{j=1}^n a_j \log \frac{a_j}{b_j} \quad (18)$$

characterizes the function $\phi(q)$ such as

- (c1) $\lim_{q \rightarrow 1} \phi(q) = 0$.
- (c2) There exists an interval (a, b) such that $a < 1 < b$, and $\phi(q)$ is differentiable on the interval $(a, 1) \cup (1, b)$.
- (c3) There exists a positive number k such that $\lim_{q \rightarrow 1} \frac{d\phi(q)}{dq} = -\frac{1}{k}$.

Proof: Eq.(18) implies condition (c1), since $\lim_{q \rightarrow 1} \left(1 - \sum_{i=1}^n a_i^q b_i^{1-q}\right) = 0$. If we differentiate $1 - \sum_{i=1}^n a_i^q b_i^{1-q}$ by q , then we obtain condition (c2). Moreover, by l'Hopital's formula, we have

$$\lim_{q \rightarrow 1} D_q(A|B) = \lim_{q \rightarrow 1} \frac{-\sum_{i=1}^n a_i^q b_i^{1-q} (\log a_i - \log b_i)}{\frac{d\phi(q)}{dq}}$$

which implies condition (c3). ■

Proposition IV.6 The condition whereby $D_q(A|U)$ takes the minimum value for a fixed posterior probability distribution as a uniform distribution $U = \{\frac{1}{n}, \dots, \frac{1}{n}\}$:

(R5) *Minimization* :

$$D_q(a_1, \dots, a_n | \frac{1}{n}, \dots, \frac{1}{n}) \geq D_q(\frac{1}{n}, \dots, \frac{1}{n} | \frac{1}{n}, \dots, \frac{1}{n})$$

implies

$$(c4) \quad \phi(q)(1-q) > 0 \text{ for } q \neq 1.$$

Proof: Since $1 - n^{q-1} \sum_{i=1}^n a_i^q$ is concave in a_i for $q > 1$ and convex in a_i for $0 \leq q < 1$, we obtain condition (c4) in order to satisfy the condition that $D_q(A|U)$ takes a minimum value. ■

As a simple example of $\phi(q)$ satisfying the conditions (c1) to (c4), we take $\phi(q) = 1 - q$ and $k = 1$, and then obtain the Tsallis relative entropy defined in Eq.(7). Also, setting $\phi(q) = 1 - 2^{1-q}$ and $k = 1$, we obtain the relative entropy of type β

$$D_q(A|B) = \frac{1 - \sum_{j=1}^n a_j^q b_j^{1-q}}{1 - 2^{1-q}},$$

which appeared in Eq. (7.2.46) of [13]. See [20], [21] for different approaches to the axiomatic characterization for the relative entropy of type β .

V. GENERALIZED FADDEEV'S AXIOM AND UNIQUENESS THEOREM FOR TSALLIS ENTROPY

A. Simplification of the uniqueness theorem for Tsallis entropy

The function $S_q(x_1, \dots, x_n)$ is assumed to be defined for the n -tuple (x_1, \dots, x_n) belonging to Δ_n and to take a value in \mathbb{R}^+ . In order to characterize the function $S_q(x_1, \dots, x_n)$, we introduce the following axiom, which is a slight generalization of the Faddeev's axiom.

Axiom V.1 (Generalized Faddeev's axiom)

- (F1) *Continuity*: The function $f_q(x) \equiv S_q(x, 1-x)$ with parameter $q \geq 0$ is continuous on the closed interval $[0, 1]$ and $f_q(x_0) > 0$ for some $x_0 \in [0, 1]$.
- (F2) *Symmetry*: For arbitrary permutation $\{x'_k\} \in \Delta_n$ of $\{x_k\} \in \Delta_n$,

$$S_q(x_1, \dots, x_n) = S_q(x'_1, \dots, x'_n). \quad (19)$$

- (F3) *Generalized additivity*: For $x_n = y + z$, $y \geq 0$ and $z > 0$,

$$S_q(x_1, \dots, x_{n-1}, y, z) = S_q(x_1, \dots, x_n) + x_n^q S_q\left(\frac{y}{x_n}, \frac{z}{x_n}\right). \quad (20)$$

Conditions (F1) and (F2) are identical to the original Faddeev's conditions, except for the addition of the parameter q . Condition (F3) is a generalization of the original Faddeev's additivity condition in the sense that condition (F3) of the present paper uses x_n^q as the factor of the second term in the right-hand side, whereas the original condition uses x_n as

this factor. Note that condition (F3) of the present paper is a simplification of condition (G3) in [8], because condition (F3) of the present paper does not require the summation on i from 1 to n . Moreover, the present axiom does not require the maximality condition (G2) in [8]. Therefore, the present axiom improves the generalized Shannon-Khinchin's axiom introduced in [8] for the characterization of Tsallis entropy. For the above generalized Faddeev's axiom, we have the following uniqueness theorem for Tsallis entropy:

Theorem V.2 Conditions (F1), (F2) and (F3) uniquely give the form of the function $S_q : \Delta_n \rightarrow \mathbb{R}^+$ such that

$$S_q(x_1, \dots, x_n) = -\lambda_q \sum_{i=1}^n x_i^q \ln_q x_i, \quad (21)$$

where λ_q is a positive constant that depends on the parameter $q \geq 0$.

Proof: For the special case of $q = 1$, the present theorem directly follows from the theorem shown in [4]. Thus, $q \neq 1$ is assumed. The proof of this theorem is similar to that presented by H. Tveberg [4]. From conditions (F2) and (F3), for any x, y, z satisfying $x, y \geq 0, z > 0$ and $x + y + z = 1$, $S_q(x, y, z)$ is expanded into separate equations:

$$\begin{aligned} S_q(x, y, z) &= S_q(x, y+z) + (y+z)^q S_q\left(\frac{y}{y+z}, \frac{z}{y+z}\right) \\ &= S_q(y, x+z) + (x+z)^q S_q\left(\frac{x}{x+z}, \frac{z}{x+z}\right). \end{aligned}$$

Therefore, we have

$$\begin{aligned} f_q(x) + (1-x)^q f_q\left(\frac{y}{1-x}\right) \\ = f_q(y) + (1-y)^q f_q\left(\frac{x}{1-y}\right) \end{aligned} \quad (22)$$

Since Eq.(22) is defined for any $0 \leq x < 1$ and $0 \leq y < 1$, by setting $x = 0$ and $y > 0$, we have

$$f_q(0) + f_q(y) = f_q(y) + (1-y)^q f_q(0).$$

Thus, we have

$$S_q(0, 1) = f_q(0) = 0. \quad (23)$$

Integrating both sides of Eq.(22) with respect to y from 0 to $1-x$, we have

$$\begin{aligned} \int_0^{1-x} f_q(x) dy + (1-x)^q \int_0^{1-x} f_q\left(\frac{y}{1-x}\right) dy \\ = \int_0^{1-x} f_q(y) dy + \int_0^{1-x} (1-y)^q f_q\left(\frac{x}{1-y}\right) dy, \end{aligned}$$

which can be deformed as follows:

$$\begin{aligned} (1-x) f_q(x) + (1-x)^{q+1} \int_0^1 f_q(t) dt \\ = \int_0^{1-x} f_q(t) dt + x^{q+1} \int_x^1 t^{-q-2} f_q(t) dt. \end{aligned} \quad (24)$$

Since the function $f_q(x)$ is continuous on the closed interval $[0, 1]$ due to condition (F1), it is differentiable on the open

interval $x \in (0, 1)$. By differentiating both sides of Eq.(24) and applying the relation

$$f_q(x) = f_q(1-x) \quad (25)$$

due to condition (F2), we have

$$(1-x)f'_q(x) = (q+1)(1-x)^q \int_0^1 f_q(t) dt + (q+1)x^q \int_x^1 t^{-q-2} f_q(t) dt - \frac{f_q(x)}{x}. \quad (26)$$

Since $f'_q(x)$ is also differentiable on $(0, 1)$, by again differentiating both sides of Eq.(26), we have

$$(1-x)f''_q(x) = -q(q+1)(1-x)^{q-1} \int_0^1 f_q(t) dt + q(q+1)x^{q-1} \int_x^1 t^{-q-2} f_q(t) dt - \frac{q f_q(x)}{x^2} - \frac{f'_q(x)}{x} + f'_q(x) \quad (27)$$

Multiplying both sides of Eq.(27) by x , we have

$$x(1-x)f''_q(x) = -q(q+1)x(1-x)^{q-1} \int_0^1 f_q(t) dt + q(q+1)x^q \int_x^1 t^{-q-2} f_q(t) dt - \frac{q f_q(x)}{x} + (x-1)f'_q(x). \quad (28)$$

Also, multiplying both sides of Eq.(26) by q , we obtain

$$q(1-x)f'_q(x) = q(q+1)(1-x)^q \int_0^1 f_q(t) dt + q(q+1)x^q \int_x^1 t^{-q-2} f_q(t) dt - \frac{q f_q(x)}{x}. \quad (29)$$

From Eq.(28) and Eq.(29), we have the following differential equation:

$$x f''_q(x) = (q-1)f'_q(x) - q\lambda_q(1-x)^{q-2}, \quad (30)$$

where $\lambda_q \equiv (q+1) \int_0^1 f_q(t) dt$. This differential equation has the following general solution with constants $c_1(q)$ and $c_2(q)$, which depend only on parameter q :

$$f_q(x) = c_1(q) + c_2(q)x^q + \frac{\lambda_q(1-x)^q}{1-q}. \quad (31)$$

The initial condition Eq.(23) implies $c_1(q) = \frac{\lambda_q}{q-1}$. In addition, the condition Eq.(25) implies $c_2(q) = \frac{\lambda_q}{1-q}$. Substituting $c_1(q)$ and $c_2(q)$ into Eq.(31), we have

$$f_q(x) = -\lambda_q \{x^q \ln_q x + (1-x)^q \ln_q(1-x)\}, \quad (32)$$

after the calculations. Since there exists some $t_0 \in [0, 1]$ such that $f_q(t_0) > 0$, due to condition (F1), and the range of S_q is \mathbb{R}^+ , we have $\lambda_q > 0$ for any $q \geq 0$. Therefore, Eq.(21) is proven for $n = 2$. Finally, Eq.(21) is proven for the general

$n \geq 3$ by induction on n . Assuming that Eq.(21) is true for any $n \in \mathbb{N}$, we obtain the following calculations:

$$\begin{aligned} S_q(x_1, \dots, x_n, x_{n+1}) &= S_q(x_1, \dots, x_n + x_{n+1}) \\ &\quad + (x_n + x_{n+1})^q S_q\left(\frac{x_n}{x_n + x_{n+1}}, \frac{x_{n+1}}{x_n + x_{n+1}}\right) \\ &= -\lambda_q \sum_{i=1}^{n-1} x_i^q \ln_q x_i - \lambda_q (x_n + x_{n+1})^q \ln_q (x_n + x_{n+1}) \\ &\quad - \lambda_q (x_n + x_{n+1})^q \left\{ \left(\frac{x_n}{x_n + x_{n+1}}\right)^q \ln_q \left(\frac{x_n}{x_n + x_{n+1}}\right) \right. \\ &\quad \left. + \left(\frac{x_{n+1}}{x_n + x_{n+1}}\right)^q \ln_q \left(\frac{x_{n+1}}{x_n + x_{n+1}}\right) \right\} \\ &= -\lambda_q \sum_{i=1}^{n-1} x_i^q \ln_q x_i - \lambda_q (x_n + x_{n+1})^q \ln_q (x_n + x_{n+1}) \\ &\quad - \lambda_q x_n^q \ln_q \left(\frac{x_n}{x_n + x_{n+1}}\right) - \lambda_q x_{n+1}^q \ln_q \left(\frac{x_{n+1}}{x_n + x_{n+1}}\right) \\ &= -\lambda_q \sum_{i=1}^{n-1} x_i^q \ln_q x_i + \lambda_q (x_n + x_{n+1}) \ln_q \frac{1}{x_n + x_{n+1}} \\ &\quad - \lambda_q x_n^q \left(\ln_q x_n + x_n^{1-q} \ln_q \frac{1}{x_n + x_{n+1}} \right) \\ &\quad - \lambda_q x_{n+1}^q \left(\ln_q x_{n+1} + x_{n+1}^{1-q} \ln_q \frac{1}{x_n + x_{n+1}} \right) \\ &= -\lambda_q \sum_{i=1}^{n+1} x_i^q \ln_q x_i. \end{aligned}$$

This shows that Eq.(21) is also true for $n+1$. Thus, the proof of this theorem is completed. ■

Remark V.3 If further conditions are imposed on Axiom V.1 such that $S_q(\frac{1}{2}, \frac{1}{2}) = 1$, which is the normalization condition to characterize the structural α -entropy in [9], then we obtain $\lambda_q = \frac{1}{\ln_q 2}$ from Eq.(32). Thus, we straightforwardly obtain the structural α -entropy [9]:

$$S_q(x_1, \dots, x_n) = \frac{1 - \sum_{i=1}^n x_i^q}{1 - 2^{1-q}},$$

in a manner similar to the induction of the proof in Theorem V.2. This means that the present theorem includes the Havrda-Charvát's axiom as a special case. Note that the present axiom requires the symmetry condition, whereas the Havrda-Charvát's axiom requires the expandability condition.

B. A relation to the generalized Shannon-Khinchin's axiom

Finally, the relationship between the generalized Shannon-Khinchin's axiom introduced in [8] and the generalized Fadd-deev's axiom is studied.

Proposition V.4 Axiom III.1 implies Axiom V.1.

Proof: The fact that conditions (G1) and (G2) imply condition (F1) is trivial, thus we show conditions (G1) and (G3) imply condition (F2). If all x_i , ($i = 1, \dots, n$) are positive rational numbers, then each x_i can be represented by $\frac{l_i}{m}$,

($2 \leq l_i \leq m$, $l_i, m \in \mathbb{Z}$). Applying condition (G3), since $x_i = \frac{l_i}{m} = \sum_{j=1}^{l_i} \frac{1}{m}$, we have

$$\begin{aligned} S_q(x_1, \dots, x_n) &= S_q\left(\frac{l_1}{m}, \dots, \frac{l_n}{m}\right) \\ &= S_q\left(\underbrace{\frac{1}{m}, \dots, \frac{1}{m}}_{l_1}, \dots, \underbrace{\frac{1}{m}, \dots, \frac{1}{m}}_{l_n}\right) \\ &\quad - \sum_{i=1}^m x_i^q S_q\left(\frac{1}{l_i}, \dots, \frac{1}{l_i}\right). \end{aligned}$$

The first term in the right-hand side of the above equation does not depend on the order of (l_1, \dots, l_n) . In addition, the method how to take the summation in the second term of the right-hand side of the above equation is arbitrary, so that the above equation is equivalent to

$$S_q\left(\underbrace{\frac{1}{m}, \dots, \frac{1}{m}}_{l'_1}, \dots, \underbrace{\frac{1}{m}, \dots, \frac{1}{m}}_{l'_n}\right) - \sum_{i=1}^m x_i'^q S_q\left(\frac{1}{l'_i}, \dots, \frac{1}{l'_i}\right)$$

for the permutation $\{x_i'\}$ from $\{x_i\}$ where $x_i' = \frac{l'_i}{m}$, ($2 \leq l'_i \leq m$, $l'_i, m \in \mathbb{Z}$). That is, condition (F2) holds for any rational numbers x_i . If x_i is not a rational number, then the continuity of condition (G1) after the approximation of x_i by a rational number is used, and then we have condition (F2). Finally, conditions (G3) and (G4) are applied to imply (F3). From conditions (G3), (G4) and (F2), we have

$$\begin{aligned} S_q\left(\frac{1}{2}, \frac{1}{2}\right) &= S_q\left(\frac{1}{2}, \frac{1}{2}, 0, 0\right) = S_q\left(\frac{1}{2}, 0, \frac{1}{2}, 0\right) \\ &= S_q\left(\frac{1}{2}, \frac{1}{2}\right) + \frac{1}{2^q} S_q(1, 0) + \frac{1}{2^q} S_q(1, 0). \end{aligned}$$

We therefore obtain $S_q(1, 0) = 0$, and thus it follows that

$$\begin{aligned} S_q(x_1, \dots, x_{n-1}, y, z) &= S_q(x_1, 0, x_2, 0, \dots, x_{n-1}, 0, y, z) \\ &= S_q(x_1, \dots, x_n) + \sum_{i=1}^{n-1} x_i^q S_q(1, 0) + x_n^q S_q\left(\frac{y}{x_n}, \frac{z}{x_n}\right) \\ &= S_q(x_1, \dots, x_n) + x_n^q S_q\left(\frac{y}{x_n}, \frac{z}{x_n}\right), \end{aligned}$$

which implies condition (F3). ■

In addition, we have the following proposition:

Proposition V.5 $S_q(X) = -\lambda_q \sum_{i=1}^n x_i^q \ln_q x_i$ satisfies Axiom III.1.

Proof: Conditions (G1) and (G4) are trivial. Condition (G2) is proven using the non-negativity of the Tsallis relative entropy:

$$D_q(X|Y) \equiv - \sum_{i=1}^n x_i \ln_q \frac{y_i}{x_i}$$

for two random variables X and Y , where $\{x_i\}$ and $\{y_i\}$, ($i = 1, \dots, n$) are probability distributions of X and Y ,

respectively. See [19] for the mathematical properties of the Tsallis relative entropy. Its non-negativity can be easily proven by the convexity of the function $-\ln_q(x)$. The non-negativity implies $S_q(X) \leq \ln_q n$ by setting a random variable $U = \{\frac{1}{n}, \dots, \frac{1}{n}\}$ having a uniform distribution, as a special case of Y . The maximum value is attained when $X = \{\frac{1}{n}, \dots, \frac{1}{n}\}$. Note that λ_q does not depend on the method used to take the maximum of $S_q(X)$. Thus, condition (G2) is proven. Finally, condition (G3) is proven by direct calculation.

$$\begin{aligned} S_q(x_1, \dots, x_n) &+ \sum_{i=1}^n x_i^q S_q\left(\frac{x_{i1}}{x_i}, \dots, \frac{x_{im_i}}{x_i}\right) \\ &= -\lambda_q \sum_{i=1}^n x_i^q \ln_q x_i + x_1^q S_q\left(\frac{x_{11}}{x_1}, \dots, \frac{x_{1m_1}}{x_1}\right) \\ &\quad + \dots + x_n^q S_q\left(\frac{x_{n1}}{x_n}, \dots, \frac{x_{nm_n}}{x_n}\right) \\ &= -\lambda_q \sum_{i=1}^n x_i^q \ln_q x_i \\ &\quad - \lambda_q x_1^q \left\{ \left(\frac{x_{11}}{x_1}\right)^q \ln_q \frac{x_{11}}{x_1} + \dots + \left(\frac{x_{1m_1}}{x_1}\right)^q \ln_q \frac{x_{1m_1}}{x_1} \right\} \\ &\quad - \dots \\ &\quad - \lambda_q x_n^q \left\{ \left(\frac{x_{n1}}{x_n}\right)^q \ln_q \frac{x_{n1}}{x_n} + \dots + \left(\frac{x_{nm_n}}{x_n}\right)^q \ln_q \frac{x_{nm_n}}{x_n} \right\} \\ &= -\lambda_q \sum_{i=1}^n x_i^q \ln_q x_i - \lambda_q \left(x_{11}^q \ln_q x_{11} + x_{11} \ln_q \frac{1}{x_1} + \dots \right. \\ &\quad \left. + x_{1m_1}^q \ln_q x_{1m_1} + x_{1m_1} \ln_q \frac{1}{x_1} \right) - \dots \\ &\quad - \lambda_q \left(x_{n1}^q \ln_q x_{n1} + x_{n1} \ln_q \frac{1}{x_n} + \dots \right. \\ &\quad \left. + x_{nm_1}^q \ln_q x_{nm_1} + x_{nm_1} \ln_q \frac{1}{x_n} \right) \\ &= \lambda_q \left(x_1 \ln_q \frac{1}{x_1} + \dots + x_n \ln_q \frac{1}{x_n} \right) \\ &\quad + S_q(x_{11}, \dots, x_{1m_1}, \dots, x_{n1}, \dots, x_{nm_n}) \\ &\quad - \lambda_q \left\{ (x_{11} + \dots + x_{1m_1}) \ln_q \frac{1}{x_1} + \dots \right. \\ &\quad \left. + (x_{n1} + \dots + x_{nm_n}) \ln_q \frac{1}{x_n} \right\} \\ &= S_q(x_{11}, \dots, x_{1m_1}, \dots, x_{n1}, \dots, x_{nm_n}). \end{aligned}$$

From Theorem V.2, Proposition V.4 and Proposition V.5, we have the following equivalent relation among Axiom V.1, Axiom III.1 and Tsallis entropy: ■

Theorem V.6 The following three statements are equivalent.

- (1) $S_q : \Delta_n \rightarrow \mathbb{R}^+$ satisfies Axiom III.1.
- (2) $S_q : \Delta_n \rightarrow \mathbb{R}^+$ satisfies Axiom V.1.
- (3) For $(x_1, \dots, x_n) \in \Delta_n$, there exists $\lambda_q > 0$ such that

$$S_q(x_1, \dots, x_n) = -\lambda_q \sum_{i=1}^n x_i^q \ln_q x_i.$$

VI. CONCLUSIONS

The uniqueness theorem for Tsallis entropy introduced by Suyari [8] was generalized to the case of Tsallis relative entropy and was simplified according to the manner of Faddeev [2], [4].

Tsallis relative entropy was characterized by the generalized Hobson's axiom. The present result includes the uniqueness theorem proven by Suyari as a special case, in the sense that the choice of a trivial distribution for $B = \{b_j\}$ of Tsallis relative entropy produces the essential form of Tsallis entropy. However, note that the present theorem requires the symmetry of condition (R2), whereas that of Suyari does not.

Moreover, Tsallis entropy was characterized by the generalized Faddeev's axiom, which is a simplification of the generalized Shannon-Khinchin's axiom introduced in [8], and the uniqueness theorem proved in [8] was slightly improved by introducing the generalized Faddeev's axiom. At the same time, the present result provides a generalization of the uniqueness theorem for Shannon entropy by means of the Faddeev's axiom.

ACKNOWLEDGEMENT

The author would like to thank Professor H. Suyari for providing the opportunity to read his paper [8] before publication. The author would also like to thank the reviewers for providing valuable comments.

REFERENCES

- [1] A.I.Khinchin, The concept of entropy in probability theory, Uspekhi Mat. Nauk, Vol.8, pp.3-20(1953), (in Russian).
- [2] A.D.Faddeev, On the notion of entropy of a finite probability space, Uspekhi Mat. Nauk, Vol.11, pp.227-231(1956), (in Russian).
- [3] A.I.Khinchin, Mathematical foundations of information theory, Dover Publications, Inc., New York, 1957.
- [4] H.Tveberg, A new derivation of the information function, Math. Scand., Vol.6, pp.297-298(1958).
- [5] A.Feinstein, Foundations of information theory, McGraw-Hill Book, Co., New York, 1958.
- [6] Y.Kakihara, Abstract methods in information theory, World Scientific, 1999.
- [7] A.Hobson, A new theorem of information theory, J.Stat.Phys., Vol.1, pp.383-391(1969).
- [8] H.Suyari, Generalization of Shannon-Khinchin axioms to nonextensive systems and the uniqueness theorem for the nonextensive entropy, IEEE Trans. Inform. Theory, Vol.50, pp.1783-1787 (2004).
- [9] J.Havrdá and F.Charvát, Quantification method of classification processes, Concept of structural α -entropy, Kybernetika(Prague), Vol.3, pp.30-35(1967).
- [10] A.Rényi, On measures of entropy and information, Proc.Forth Berkeley Symp.Math.Stat.Prob., Vol.1, pp.547-561, Univ.California Press, Berkeley, 1961.
- [11] A.Rényi, Wahrscheinlichkeitsrechnung. Mit einem Anhang über Informationstheorie., VEB Deutscher Verlag der Wissenschaften, Berlin, 1962.
- [12] Z.Daróczy, General information functions, Information and Control, Vol.16, pp.36-51(1970).
- [13] J.Aczél and Z.Daróczy, On measures of information and their characterizations, Academic Press, 1975.
- [14] C.Tsallis, Possible generalization of Boltzmann-Gibbs statistics, J.Stat.Phys., Vol.52, pp.479-487(1988).
- [15] C. Tsallis et al., Nonextensive Statistical Mechanics and Its Applications, edited by S. Abe and Y. Okamoto (Springer-Verlag, Heidelberg, 2001); see also the comprehensive list of references at <http://tsallis.cat.cbpf.br/biblio.htm>.
- [16] M.Shiino, H-theorem with generalized relative entropies and the Tsallis statistics, J.Phys.Soc.Japan, Vol.67, pp.3658-3660(1998).

- [17] C.Tsallis, Generalized entropy-based criterion for consistent testing, Phys.Rev.E, Vol.58, pp.1442-1445(1998).
- [18] A.K.Rajagopal and S.Abe, Implications of form invariance to the structure of nonextensive entropies, Phys. Rev. Lett., Vol.83, pp.1711-1714(1999).
- [19] S.Furuichi, K.Yanagi and K.Kuriyama, Fundamental properties of Tsallis relative entropy, J.Math.Phys., Vol.45, No.12, pp. 4868-4877 (2004).
- [20] P.N.Rathie and P.Kannappan, A directed-divergence function of type β , Information and Control, Vol.20, pp.38-45 (1972).
- [21] N.Muraki and T.Kawaguchi, On a generalization of Havrdá-Charvát's α -entropy to relative α -entropy and its properties in a continuous system, TENSOR, Vol.46, pp.154-167 (1987).

PLACE
PHOTO
HERE

Shigeru FURUICHI received the B.S. degree from Department of Mathematics, Tokyo University of Science, and M.S. and Ph.D. from Department of information Science, Tokyo University of Science, Japan. He was an Assistant Professor during 1997-2001 and has been a Lecturer from 2001 in Department of Electronics and Computer Science, Tokyo University of Science, Yamaguchi, Japan. His research interests include information theory, entropy theory and operator theory.